

AD-A157 296 BAYESIAN EXTENSIONS TO A BASIC MODEL OF SOFTWARE
RELIABILITY(U) CALIFORNIA UNIV BERKELEY OPERATIONS
RESEARCH CENTER W S JEWELL JUN 85 ORC-85-4
UNCLASSIFIED AFOSR-81-0122 F/G 9/2

BAYESIAN EXTENSIONS TO A BASIC MODEL OF SOFTWARE
RELIABILITY(U) CALIFORNIA UNIV BERKELEY OPERATIONS
RESEARCH CENTER W S JEWELL JUN 85 ORC-85-4
AFOSR-81-0122 F/G 9/

1/1

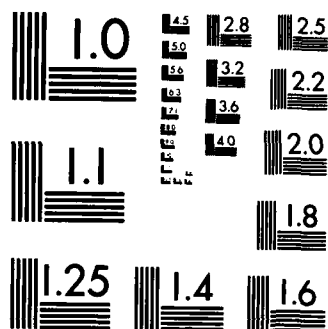
UNCLASSIFIED

F/G 9/2

NL

END

492 *Journal of Management Education* 36(4)



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A157 296

OPERATIONS RESEARCH CENTER

BAYESIAN EXTENSIONS TO A BASIC MODEL
OF SOFTWARE RELIABILITY

by

William S. Jewell[†]

ORC-85-4

June 1985

DTIC FILE COPY

UNIVERSITY OF CALIFORNIA



BERKELEY

DTIC
ELECTE
JUL 19 1985
S D
G

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

85 7 10 024

BAYESIAN EXTENSIONS TO A BASIC MODEL
OF SOFTWARE RELIABILITY

by

William S. Jewell[†]

✓
ORC- 85-4

June 1985

DTIC
ELECTE
S JUL 1 9 1985 D
G

[†]Department of Industrial Engineering and Operations Research,
University of California, Berkeley, California.

This research was supported by the Air Force Office of Scientific
Research (AFSC), USAF, under Grant AFOSR-81-0122 with the University
of California. Reproduction in whole or in part is permitted for
any purpose of the United States Government.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

| | | | | | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|------------------------------------|------------------------------|
| 1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED | | | 1b. RESTRICTIVE MARKINGS | | |
| 2a. SECURITY CLASSIFICATION AUTHORITY | | | 3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited. | | |
| 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE | | | | | |
| 4. PERFORMING ORGANIZATION REPORT NUMBER(S) ORC-85-4 | | | 5. MONITORING ORGANIZATION REPORT NUMBER(S) | | |
| 6a. NAME OF PERFORMING ORGANIZATION Operations Research Center | | 6b. OFFICE SYMBOL (If applicable) | 7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research | | |
| 6c. ADDRESS (City, State and ZIP Code) University of California Berkeley, California 94720 | | | 7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448 | | |
| 8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR | | 8b. OFFICE SYMBOL (If applicable) NM | 9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-81-0122 | | |
| 8c. ADDRESS (City, State and ZIP Code) Bolling AFB DC 20332-6448 | | | 10. SOURCE OF FUNDING NOS. | | |
| | | | PROGRAM ELEMENT NO. 61102F | PROJECT NO. 2304 | TASK NO. |
| | | | WORK UNIT NO. | | |
| 11. TITLE (Include Security Classification) BAYESIAN EXTENSIONS TO A BASIC MODEL OF SOFTWARE RELIABILITY | | | | | |
| 12. PERSONAL AUTHOR(S) William S. Jewell | | | | | |
| 13a. TYPE OF REPORT Technical Report | | 13b. TIME COVERED FROM _____ TO _____ | | 14. DATE OF REPORT (Yr., Mo., Day) | |
| | | | | 15. PAGE COUNT 24 | |
| 16. SUPPLEMENTARY NOTATION | | | | | |
| 17. COSATI CODES | | | 18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) | | |
| FIELD | GROUP | SUB. GR. | Keywords: software reliability, defect removal, Bayesian prediction | | |
| | | | | | |
| | | | | | |
| 19. ABSTRACT (Continue on reverse if necessary and identify by block number) (See Abstract) | | | | | |
| 20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/> | | | 21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED | | |
| 22a. NAME OF RESPONSIBLE INDIVIDUAL Major Brian W. Woodruff | | | 22b. TELEPHONE NUMBER (Include Area Code) (202) 767-5027 | | 22c. OFFICE SYMBOL NM |

- 2 -

ABSTRACT

✓ A Bayesian analysis of the software reliability model of Jelinski and Moranda is given, based upon Meinhold and Singpurwalla. Important extensions are provided to the stopping rule and prior distribution of the number of defects, as well as permitting uncertainty in the failure rate. It is easy to calculate the predictive distribution of unfound errors at the end of software testing, and to see the relative effects of uncertainty in the number of errors and in the detection efficiency. The behavior of the predictive mode and mean over time are examined as possible point estimators, but are clearly inferior to calculating the full predictive distribution. ←

| | |
|--------------------|--------------------------------------------|
| Accession For | |
| NTIS GRA&I | <input checked="checked" type="checkbox"/> |
| DTIC TAB | <input type="checkbox"/> |
| Unannounced | <input type="checkbox"/> |
| Justification | |
| By _____ | |
| Distribution/ | |
| Availability Codes | |
| Dist | Avail and/or Special |
| A/1 | |



BAYESIANS EXTENSIONS TO A BASIC MODEL OF SOFTWARE RELIABILITY

by

William S. Jewell

1. INTRODUCTION

A convenient method of surveying the problems of concern in software reliability is to scan the articles in the August, 1979 issue of the IEEE Transactions on Reliability (Volume R-28:3). In addition to various reliability management and computer science issues, it refers to the process of identifying and removing computer software errors through program testing.

One of the earliest and simplest models for describing the stochastic process of error occurrence is that proposed by Jelinski and Moranda [2]. Essentially, they assume that a fixed, but unknown number of defects or "bugs" in the program are "competing" equally and randomly for discovery, each at the same average, unknown rate. The objective of the analysis is to estimate the number of undetected faults remaining in the program after a certain time and/or failure history, or to make other performance guarantees (Littlewood, [9]).

Parenthetically, we should note that many variations to this basic model have been proposed--Littlewood and Verrall [7] [8], Goel and Okumoto [1], and Langberg and Singpurwalla [6] analyze those competing models and show that they can be cast in a unified framework. More recent elaborations are in Moranda [12], Littlewood [10], and Kremer [5].

In this paper we shall stick with the basic Jelinski-Moranda (J-M) [2] model for simplicity, and to highlight our Bayesian extensions. At first, our paper is tutorial, following the Bayesian development of Meinhold and Singpurwalla (M-S) [11]. However, the following important extensions are made:

- (a) The testing protocol is permitted to run for a fixed length of time--possibly, but not necessarily, coinciding with a failure epoch;
- (b) The distribution of the unknown number of defects is generalized from the one-parameter Poisson distribution by assuming that the parameter is itself a random quantity with a Beta prior distribution;
- (c) Although the calculation of the posterior distributions of the parameters leads to complex expressions, we show that the computation of the predictive distribution for undetected errors is straightforward;
- (d) Although it is now recognized that the MLE's for reliability growth can be very unstable, we show that, if a point estimator is needed, the predictive mode is easily calculated without obtaining the full distribution first.

The paper concludes with a numerical example.

2. BASIC MODEL

Suppose there are N defects in a given software product and that, as a result of exercising the program, failures are observed to occur at epochs $t_{(1)} < t_{(2)} < \dots$. Each failure initiates a debugging action that permanently eliminates the error causing the failure; for simplicity we assume that the time to find and fix the error is negligible, or, what is the same thing, we measure only the duration of time while running the program. If we then assume:

- (a) The overall failure rate of the program, $\rho(t)$ hours⁻¹, between $t_{(i)}$ and $t_{(i+1)}$, is a constant equal to $(N-i)\emptyset$ hours⁻¹, ($i = 0, 1, 2, \dots, N-1$; $t_{(0)} = 0$), where \emptyset is a unit failure rate; and,
- (b) given N and \emptyset , the inter-failure times, $\tau_i = t_{(i)} - t_{(i+1)}$ ($i = 0, 1, \dots, N-1$), are statistically independent;

then it follows that the interfailure times have a state-dependent exponential distribution of the form:

$$\Pr\{\tilde{\tau}_i > \tau\} = \exp[-\emptyset(N-i+1)\tau]. \quad (2.1)$$

The resulting point process, a typical realization of which is shown in Figure 1, is a pure-death continuous-time Markov process. (The total failure rate is a random variable not only because of the random locations of the failure epochs but also because \emptyset and N can be considered as random quantities.)

This process can also be viewed from a competing-risks point of view, in which each error has an independent probability, $\emptyset dt$, of causing a failure in $(t, t + dt)$. The individual "lifetimes" until an error is uncovered, $\{t_i; (i = 1, 2, \dots, N)\}$ are then mutually independent and exponentially distributed with parameter \emptyset . Our assumptions above then mean that

all N defects are considered to be "on test" simultaneously until failure. After an interval $(0, t]$, we will observe that a certain random number, $\tilde{n}(t) = n(t)$, of them will have failed with observed lifetimes $\{t_1, t_2, \dots, t_{n(t)}\}$, and $N - n(t)$ of them will still be "a failure waiting to occur." Since the probability is $Q = \exp(-\emptyset t)$ that a bug will *not* surface in $(0, t]$, and $P = 1 - Q$ that it will, it follows that the distribution of $\tilde{n}(t)$, given N and \emptyset , is Binomial (N, P) :

$$\Pr\{\tilde{n}(t) = n\} = \binom{N}{n} (1 - e^{-\emptyset t})^n e^{-\emptyset(N-n)}, \quad (n = 0, 1, \dots, N) \quad (2.2)$$

with moments:

$$E\{\tilde{n}(t)\} = N(1 - e^{-\emptyset t}) ; \quad V\{\tilde{n}(t)\} = Ne^{-\emptyset t}(1 - e^{-\emptyset t}) . \quad (2.3)$$

(These results also follow from the solution of the birth-and-death equations of the Markov process description.)

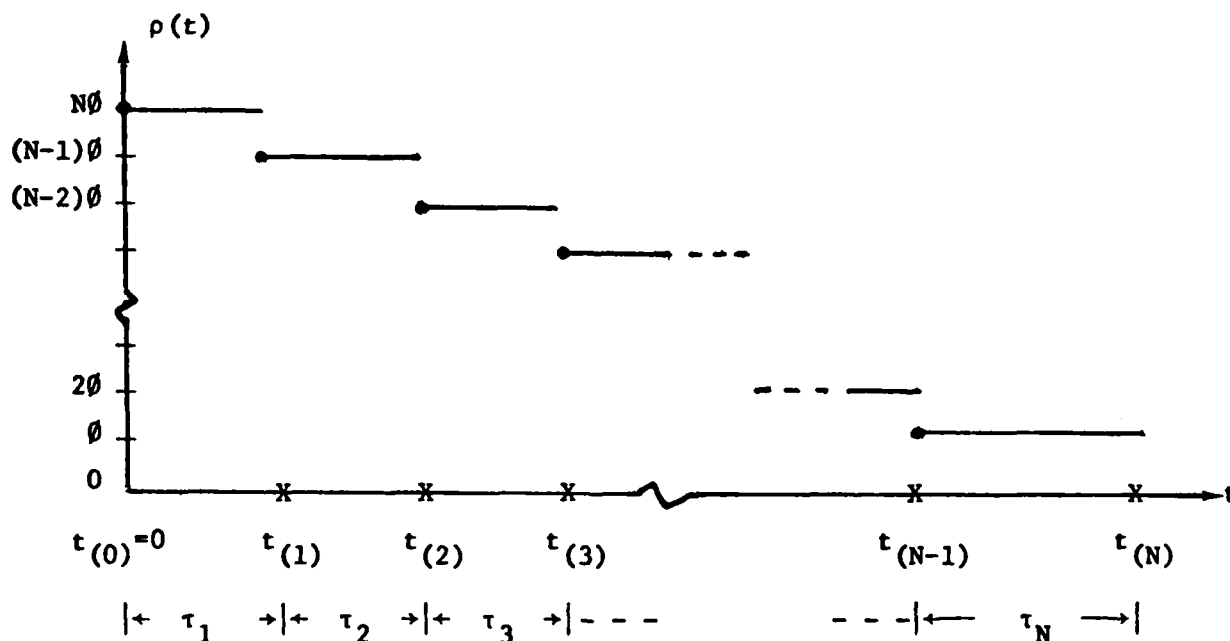


Figure 1. Typical realization of overall failure rate over time

From (2.2) we can in principle find the distribution of $\tilde{\rho}(t)$ at any t ; in particular, the *mean failure rate* is the function

$$E\{\tilde{\rho}(t)\} = (N - n(t))\theta = N\theta e^{-\theta t} \quad (2.4)$$

if both θ and N are known, and it is this form which suggests defect identification and removal are very much like *reliability growth models*, a point already made by early authors.

However, the software reliability problem is different in the following senses:

- (a) θ , which reflects both the rate at which certain portions of the program are exercised as well as the probability that a certain bug will cause a failure to occur (plus possibly the probability that an error will be recognized by the operator), is usually unknown a priori and has to be estimated from past debugging experience and from the failure data of this experiment;
- (b) \tilde{N} , the total number of defects in the program, is always unknown a priori;
- (c) Since all failures will ultimately be found, the estimation problem of primary interest after a testing interval of t hours is the prediction of the *distribution of unfound errors*, $\tilde{n}_0(t) = \tilde{N} - n(t)$; other possible measures of interest are estimators of the *local failure rate*, $\tilde{\rho}(t)$, and of the *remaining time until total debugging*, $t_{(\tilde{N})} - t$.

By casting this problem in a Bayesian framework, similar to L-V and M-S, we will see that these distributions can be found quite easily. In common with M-S, we shall make the reasonable assumption that \tilde{N} is Poisson distributed, with parameter $\lambda \text{ hour}^{-1}$. Since this only gives a one-parameter family, and since it is rare that λ would be known a priori, we shall enrich this

assumption by additionally assuming that $\tilde{\lambda}$ is a random quantity with a Gamma prior; this is equivalent to assuming that \tilde{N} is a priori from a Pascal (Negative Binomial) density, and permits incorporating large variances in our prior estimates. In common with L-V and M-S, we also assume that $\tilde{\theta}$ is an independent random quantity, also Gamma distributed.

3. FULL-DATA PRIORS AND POSTERIORIS

In the usual case, the full history of all failure times will be recorded, so that the data from an experiment run for t hours will be $\mathcal{D}_t = \{t_1, t_2, \dots, t_n; n\}$ where by n we mean $n(t)$. For convenience, we also switch to the unordered individual lifetimes of the n out of N failures that appear in $(0, t]$. It then follows from first principles that, given N and \emptyset , the data likelihood is:

$$\begin{aligned} p(\mathcal{D}_t \mid N, \emptyset) &= \frac{N!}{(N-n)!} \left(\prod_{i=1}^n \emptyset e^{-\emptyset t_i} \right) \left(\prod_{j=1}^{N-n} e^{-\emptyset t} \right) \\ &= \frac{N!}{(N-n)!} \emptyset^n e^{-\emptyset T}, \end{aligned} \quad (3.1)$$

where

$$T = T(\mathcal{D}_t, N) = \sum_{i=1}^n t_i + (N-n)t \quad (3.2)$$

is the familiar *total-time-on-test statistic* for N items tested in parallel. Note that (3.1) remains the likelihood for any non-informative stopping rule; for example, instead of stopping after t hours, we could also stop after the n^{th} failure, giving $T=t_{(n)}$ (see M-S). Thus, (n, T) are the usual sufficient statistics for $\tilde{\emptyset}$ when N is known.

However, in our case, N is unknown and is Poisson (λ) a priori, so that:

$$p(\mathcal{D}_t, N \mid \lambda, \emptyset) = p(\mathcal{D}_t \mid N, \emptyset) \frac{\lambda^N e^{-\lambda}}{N!}.$$

Marginalizing out the values of $N \geq n$, we get the final data likelihood, given the parameters λ and \emptyset :

$$p(\mathcal{D}_t \mid \lambda, \emptyset) = \frac{(\lambda \emptyset)^n e^{-\emptyset S} e^{-\lambda} [1 - e^{-\emptyset t}]}{n!}, \quad (3.3)$$

where

$$S = S(\mathcal{D}_t) = \sum_{i=1}^n t_i = \sum_{i=1}^n t_{(i)} = \sum_{i=1}^n (n-i+1) \tau_i \quad (3.4)$$

is the total-time-on-test for the discovered errors. Thus, (n, S) are sufficient for $(\tilde{\lambda}, \tilde{\emptyset})$.

It remains to choose appropriate priors for $\tilde{\lambda}$ and $\tilde{\emptyset}$. Suppose $\tilde{\emptyset} = \emptyset$ was, in fact, fixed and known and $\emptyset t$ was *large*; then it can be seen from (3.3) that the Gamma density:

$$p(\lambda) = \text{Ga}(\lambda \mid a, b) = \frac{b^a \lambda^{a-1} e^{-b\lambda}}{\Gamma(a)}, \quad (3.5)$$

would be a convenient choice for the prior on $\tilde{\lambda}$, since this combination would be *closed under sampling*, that is, the posterior-to-data density, $p(\lambda \mid \mathcal{D}_t)$, would *also* be Gamma, with revised parameters:

$$a' = a+n; \quad b' = b+1. \quad (3.6)$$

As mentioned previously, this is tantamount to assuming that \tilde{N} is Pascal $(a, (b+1)^{-1})$ distributed a priori; the hyperparameters a and b can be estimated from $E\{\tilde{N}\} = (a/b)$ and $V\{\tilde{N}\} = (a(b+1)/b^2)$. (See (4.2) below.)

Conversely, if $\tilde{\lambda} = \lambda$ were fixed and all values of $\tilde{Q} = e^{-\emptyset t}$ were assumed negligible compared to unity, then we see that the convenient prior for $\tilde{\emptyset}$ would also be a Gamma density, say $\text{Ga}(\emptyset \mid c, d)$, so that, posterior-to-the-data, $p(\emptyset \mid \mathcal{D})$ would still be Gamma, but with updated hyperparameters:

$$c' = c + n; \quad d' = d + S. \quad (3.7)$$

The original values of these hyperparameters could be estimated from the relationships $E\{\tilde{\theta}\} = (c/d)$ and $V\{\tilde{\theta}\} = (c/d^2)$, or from forming some opinion about the time for a single error to be found, which has moments $E\{\tilde{\theta}^{-1}\} = (d/(c-1))$ and $V\{\tilde{\theta}^{-1}\} = (d^2/(c-1)^2(c-2))$.

Now, unfortunately, our likelihood (3.3) is more complicated, and the coupling term, $\exp\{\lambda e^{-\tilde{\theta}t}\} = \exp\{\lambda Q\}$, means that, posterior-to-the-data, $p(\lambda, \tilde{\theta} \mid \mathcal{D}_t)$ will have the two parameters dependent, even though they were a priori independent. Nevertheless, the assumption of independent Gamma priors still turns out to be the most effective one for analytic simplicity, and so, using Bayes' law, we write the joint posterior on the parameters as:

$$p(\lambda, \tilde{\theta} \mid \mathcal{D}_t) \propto e^{\lambda Q} \text{Ga}(\lambda \mid a', b') \text{Ga}(\tilde{\theta} \mid c', d') \quad (3.8)$$

where the interaction is clearly seen. Of course, as $t \rightarrow \infty$, our posterior opinions about $\tilde{\lambda}$ and $\tilde{\theta}$ will become independent again, since $Q \rightarrow 0$.

By expanding the coupling term in powers of (λQ) , one can write (3.8) in closed form as the ratio of two infinite series, as in Jewell [4]. We shall not pursue this here, but pass to the more interesting problem of predicting the unfound errors.

4. PRIOR AND POSTERIOR PREDICTION OF UNFOUND ERRORS

It is well known that the partition of a Poisson-distributed random integer \tilde{N} by means of a Binomial process leads to two independent Poisson processes. It follows that the *number of undetected errors at time t*, $\tilde{n}_0(t) = \tilde{N} - n(t)$ is Poisson (λQ) , that is:

$$p(n_0 | \lambda, \emptyset, t) = \frac{(\lambda Q)^{n_0} e^{-\lambda Q}}{n_0!} \quad (Q = e^{-\emptyset t}) \quad (4.1)$$

If \emptyset were fixed, we could use (3.5) to show that our prior prediction of $\tilde{n}_0(t)$ would be the marginal density:

$$p(n_0 | \emptyset, t) = \frac{\Gamma(a + n_0)}{\Gamma(a) n_0!} \left(\frac{b}{b+Q} \right)^a \left(\frac{Q}{b+Q} \right)^{n_0}, \quad (4.2)$$

which is Negative Binomial or Pascal $(a, Q/(b+Q))$. Thus, before debugging, our prior opinion about the errors that will remain after time t would be that $E\{\tilde{n}_0 | \emptyset, t\} = (aQ/b)$ and $V\{\tilde{n}_0 | \emptyset, t\} = (aQ/b)(1+(Q/b))$, assuming always that \emptyset is known. In particular, as $t \rightarrow 0$, $Q \rightarrow 1$, and we obtain the Pascal prior for \tilde{N} previously mentioned.

If we attempt to use the Gamma (c, d) prior on $\tilde{\emptyset}$ with (4.2) to get $p(n_0 | t)$, the unconditional *prior* prediction of remaining errors at t , we again run into analytic difficulties and must settle for an infinite series. However, the moments are no difficulty, whence we find, for example, the prior expected number of errors remaining at time t :

$$E\{\tilde{n}_0 | t\} = E\{\tilde{N} \tilde{Q}\} = \left(\frac{a}{b} \right) \left(\frac{d}{d+t} \right)^c, \quad (4.3)$$

and the prior expected failure rate at time t :

$$E\{\tilde{\rho}(t)\} = E\{\tilde{n}_0 \tilde{\emptyset}\} = \left(\frac{a}{b} \right) \frac{c d^c}{(d+t)^{c+1}}. \quad (4.4)$$

Since $p(\lambda, \emptyset \mid \mathcal{D}_t)$ and $p(n_o \mid t)$ can only be found in terms of infinite series, it might seem hopeless to attempt to compute the *predictive* distribution of \tilde{n}_o , given the data. However, comparison of (3.3) and (4.1) reveals that there is a fortuitous cancellation of the coupling term when using

$$p(n_o \mid \mathcal{D}_t) = \iint p(n_o \mid \lambda, \emptyset, t) p(\lambda, \emptyset \mid \mathcal{D}_t) d\lambda d\emptyset ,$$

and we find easily the predictive density:

$$p(n_o \mid \mathcal{D}_t) = K \frac{\Gamma(a' + n_o)}{\Gamma(a') n_o!} (b')^{-n_o} \left(\frac{d'}{d' + t n_o} \right)^{c'} , \quad (4.5)$$

where only the normalizing constant, $K = p(0 \mid \mathcal{D}_t)$, requires numerical computation.

In fact, (4.5) can be simply computed by setting $p(0 \mid \mathcal{D}_t) = 1$, using the recursion:

$$\frac{p(n_o + 1 \mid \mathcal{D}_t)}{p(n_o \mid \mathcal{D}_t)} = \left(\frac{a' + n_o}{b'} \right) \left(\frac{d' + t n_o}{d' + t + t n_o} \right)^{c'} \left(\frac{1}{n_o + 1} \right) \quad (4.6)$$

and then renormalizing. Although (4.5) is not a standard density, we see from (4.6) that, as n_o gets large, the term involving c' and d' can be ignored, so the *tail* of $p(n_o \mid \mathcal{D}_t)$ is approximately Pascal $(a+1, (b+1)^{-1})$.

Unfortunately, in contrast to (4.3), the moments of (4.5) must now be found numerically.

5. MLE AND PREDICTIVE MODE

In attempting to get a point estimator from the data, the usual idea is to find the MLEs $(\hat{N}, \hat{\theta})$ from (3.1). M-S have reported that \hat{N} can be unstable for small amounts of data, and might not even exist. In our extended model (3.3) one can find the MLEs $(\hat{\lambda}, \hat{\theta})$ from:

$$\frac{n}{\hat{\lambda}} = 1 - e^{-\hat{\theta}t} ; \quad \frac{n}{\hat{\theta}} = S + \hat{\lambda} t e^{-\hat{\theta}t} ;$$

which gives for $\hat{\theta}$ alone:

$$\frac{1}{\hat{\theta}} = \frac{S}{n} + \frac{t}{e^{\hat{\theta}t} - 1} , \quad (5.1)$$

and then $\hat{\lambda}$ by substitution, and finally $\hat{n}_0 = \hat{\lambda} - n$. These estimators should exist for all values of $t > S(\mathcal{D}_t)/n(t)$, but one can, in fact, show that (5.1) gives positive $\hat{\theta}$ only for t larger than *twice* this critical ratio! In fact, solutions for the smallest permitted values of t are very unstable numerically and cannot easily be found. About the only simple thing that can be said is that, for t large, $\hat{\theta}$ approaches $n(t)/S(\mathcal{D}_t)$, which is obvious. So much for MLEs.

However, modes also give useful point estimators, if one is needed. We find from the Pascal marginal density that the *prior* mode $\hat{n}_0 = \hat{N}$ is the smallest integer not less than the solution to:

$$N^* + 1 = \frac{a + N^*}{b + 1} \quad (5.2)$$

From (4.6), it follows that the posterior-to-data *predictive* mode is the integer not less than the solution to:

$$n_0^* + 1 = \left[\frac{a' + n_0^*}{b + 1} \right] \left[\frac{d' + t n_0^*}{d' + t + t n_0^*} \right]^{c'} \quad (5.3)$$

which can be found easily through iteration. Through arguments similar to those in Jewell [4], one can show that the first term in (5.3) is an updated estimate N' of \hat{N} , and the second term is an estimate Q' of $\tilde{Q} = e^{-\tilde{\theta}t}$, so that (5.3) reads simply $n_0^* \approx N'Q'$. Limiting cases of the hyperparameters when the priors are diffuse or degenerate give various intuitive versions of (5.3). (See [4].)

Note that (5.3) exists for all $t \geq 0$ and all data. However, for small t , n_0^* may be substantially less than $E\{\tilde{n}_0 \mid \mathcal{D}_t\}$ because of the long tails of the distribution.

6. COMPARISON WITH MEINHOLD AND SINGPURWALLA'S MODEL

Apart from the generalized stopping rule, our main difference with M-S is that they assume λ is known, rather than being a random quantity that is Gamma (a, b) , a priori. The effect of this can be most easily seen in the term $(a' + n_0)/b'$ in (4.6) and the similar term in (5.3). The M-S model can be obtained by setting $a = b\lambda$, and letting $b \rightarrow \infty$; then these terms are replaced simply by λ , and no longer depend upon the observations or the values of n_0 . At the other extreme, our model can represent diffuse price knowledge about $\tilde{\lambda}$ by letting $b \rightarrow 0$ ($V\{\tilde{\lambda}\} \rightarrow \infty$); then these terms become simply $n_0 + n$ (or $n_0^* + n$), and are essentially independent of the prior. In the general use of the Gamma prior, we obtain a "credibility" mixture between $E\{\tilde{\lambda}\}$ and $n_0 + n$ for this term.

7. EXAMPLE

Suppose we set $a = 1$ and $b = 0.1$, which corresponds to:

$$E\{\tilde{\lambda}\} = E\{\tilde{N}\} = 10 ; \quad V\{\tilde{\lambda}\} = 100 ; \quad V\{\tilde{N}\} = 110 ;$$

and then set $c = 5$ and $d = 80$, that is:

$$E\{\tilde{\theta}\} = 0.0625 \quad V\{\tilde{\theta}\} = 0.000781 \quad E\{\tilde{\theta}^{-1}\} = 20 ; \quad V\{\tilde{\theta}^{-1}\} = 133.33 .$$

Picking $N = 20$ and $\theta = 0.025$, one realization of ordered failure times $\{t_{(k)} (k = 1, 2, \dots, 20)\}$ gave:

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 3.282 | 3.889 | 5.900 | 5.903 | 5.996 | 6.055 | 9.829 | 10.40 | 10.95 | 11.90 |
| 11.96 | 14.40 | 19.98 | 20.23 | 20.38 | 22.91 | 30.05 | 35.35 | 35.69 | 59.44 |

Figure 2 shows the integral solutions n_o^* and $N^* = n(t) + n_o^*$ to the predictive mode equation (5.3), versus t . For $t < t_{(1)}$, both n_o^* and N^* are zero because no failures have occurred, even though our prior expectation is $E\{\tilde{N}\} = 10$. Thereafter, the curve jumps upward at each failure epoch, followed by integer decreases as t increases without further events (Figure 2 shows only a straight-line approximation to the true behavior).

Figure 3 shows the exact predictive means $E\{\tilde{n}_o \mid \mathcal{D}_t\}$ and $E\{\tilde{N} \mid \mathcal{D}_t\}$ for $t = 0(5)80$; these were computed numerically from the full distributions. In general, these behave more smoothly than the posterior modes and seem to converge towards the true N for lower values of t . Further, for small t they are influenced by the prior, rather than starting at zero. (Figure 3 shows only a straight-line approximation to the true behavior.)

However, the main result is not these point estimators, but the fact that the full predictive density can be obtained from (4.6). Figure 2 shows the (Pascal) density obtained at the start of the experiment ($t = 0$); for our

parameters the initial mode is zero. After a couple of errors are found, the mode shifts to a positive number and the variance begins to reduce, giving curves similar to Figure 5, where $t = 10$ and $n(10) = 7$. Thereafter, the curve essentially remains unimodal, with nodes and means given by Figures 2 and 3, with the variance decreasing almost continuously. After about $t = 46$, $n(46) = 19$, the predictive density again has its mode at the origin, with mass on only a few values of n_0 ; Figure 6 shows a typical density, at $t = 70$, 10 time units after the last error has actually been found.

In this way, we see how easily one obtains full predictive information about undetected errors for use in decision problems or in making performance guarantees.

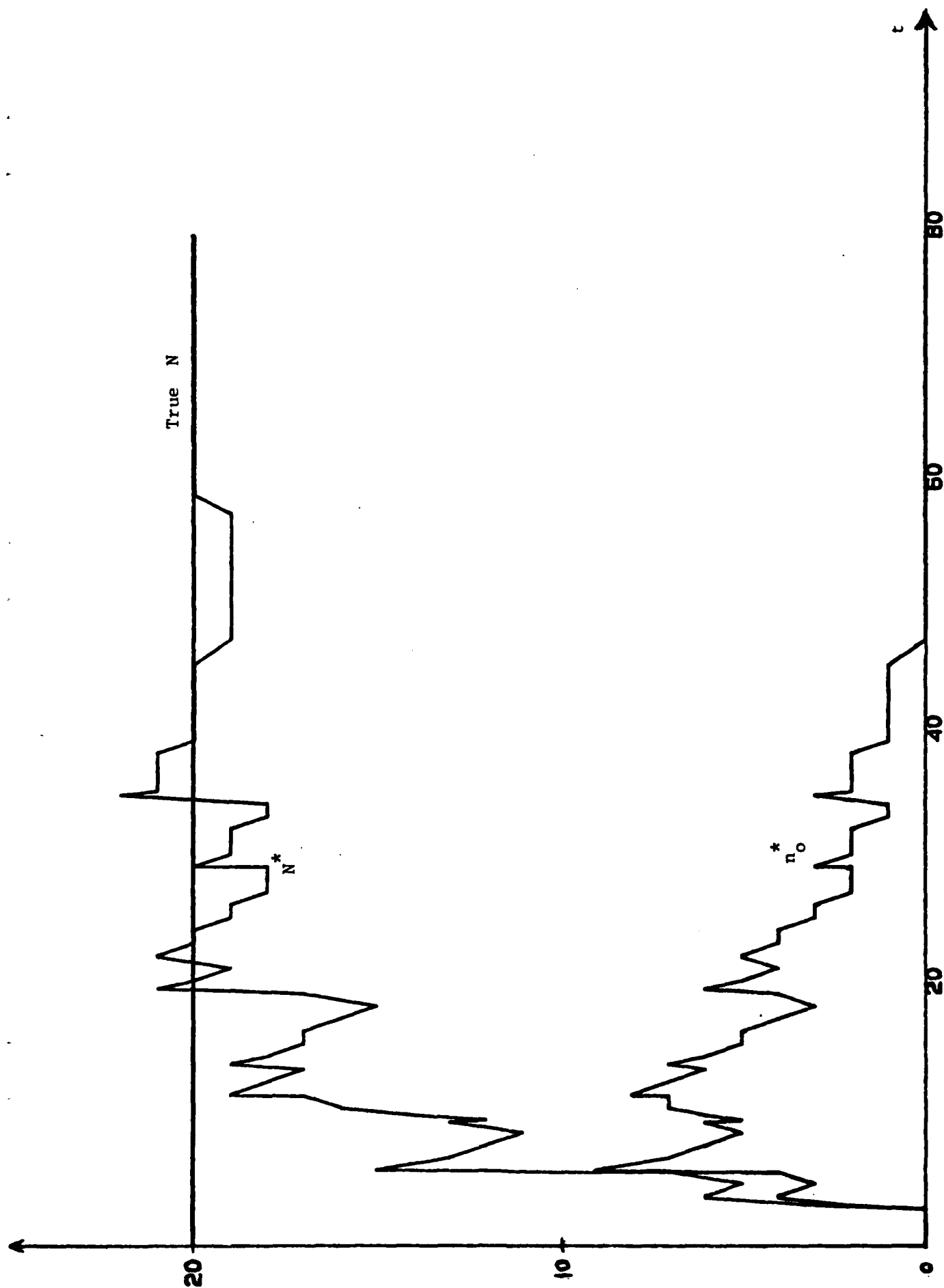


Figure 2. Approximate behavior of predictive modes, n_o^* and N^* , versus testing interval

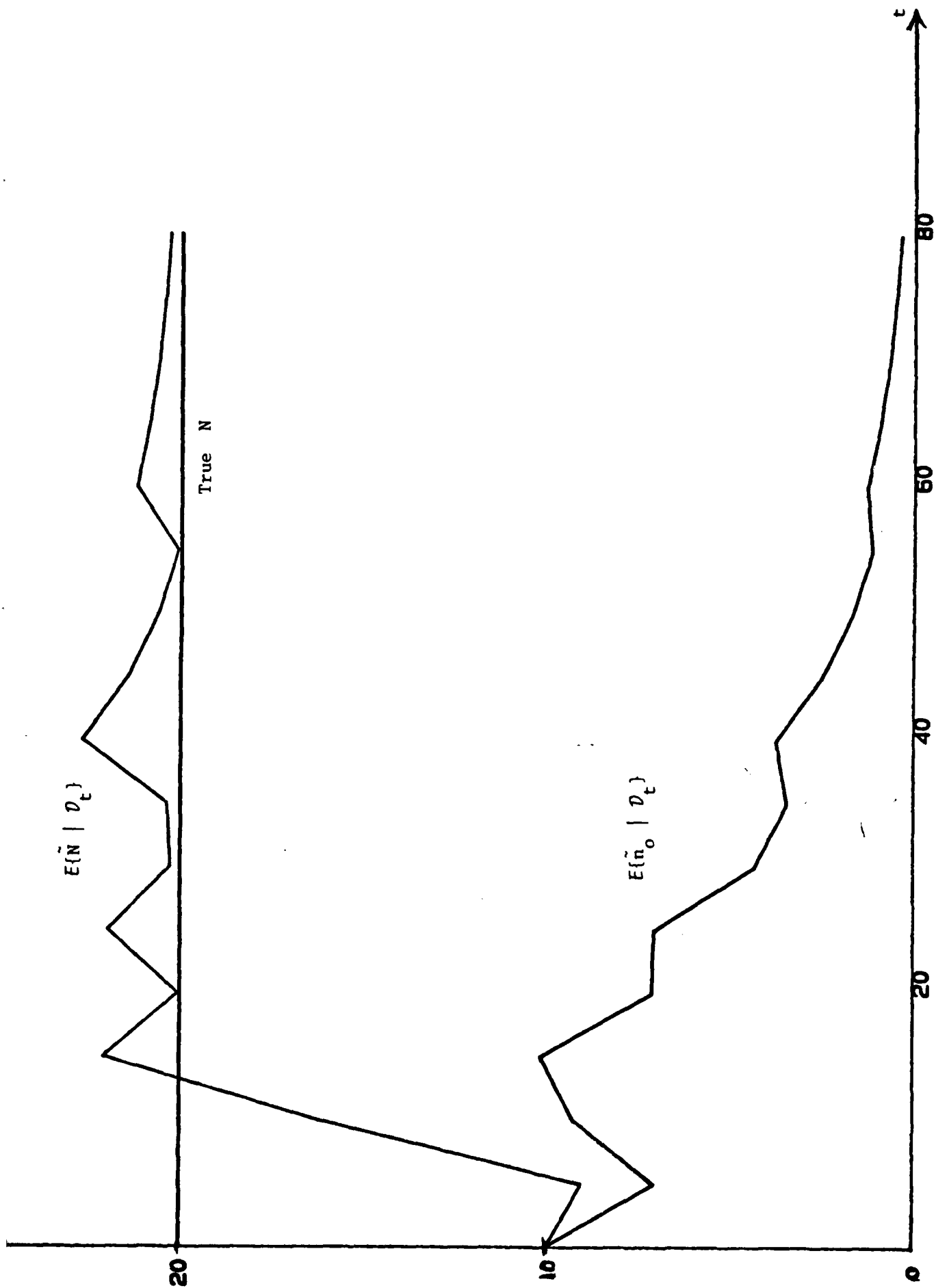


Figure 3. Approximate behavior of predictive means versus testing interval

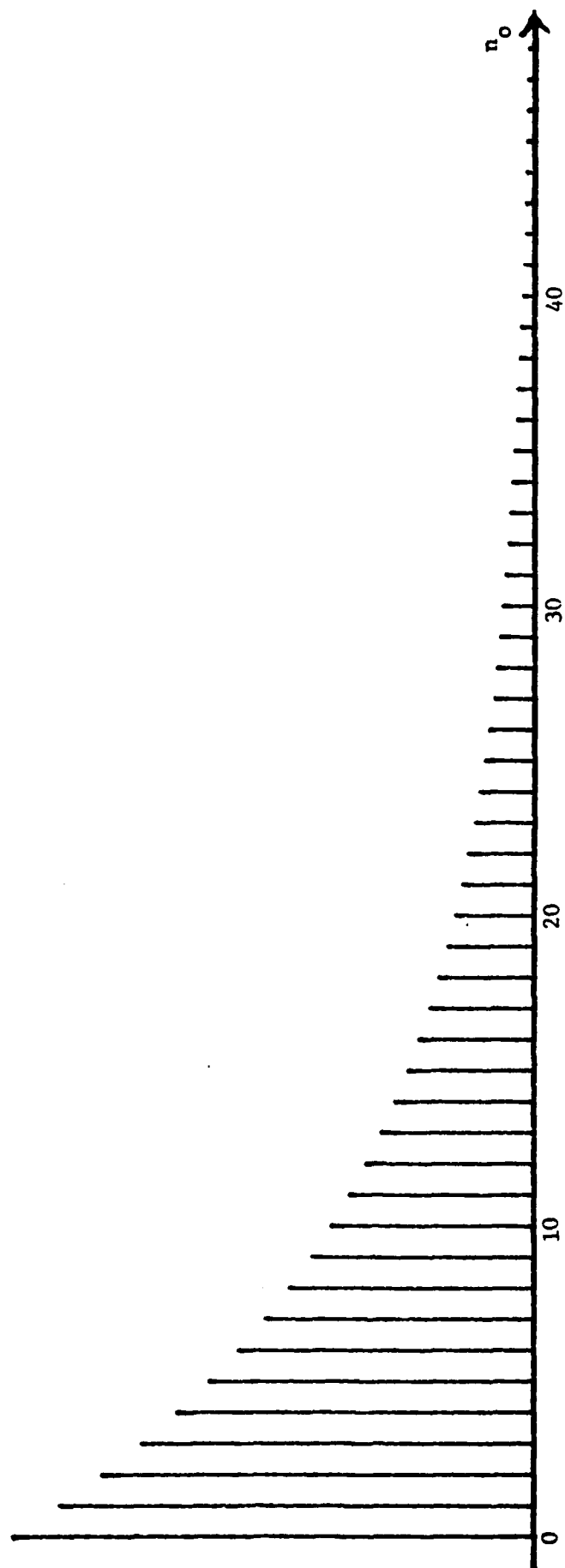


Figure 4. Predictive density of undetected errors at $t = 0$

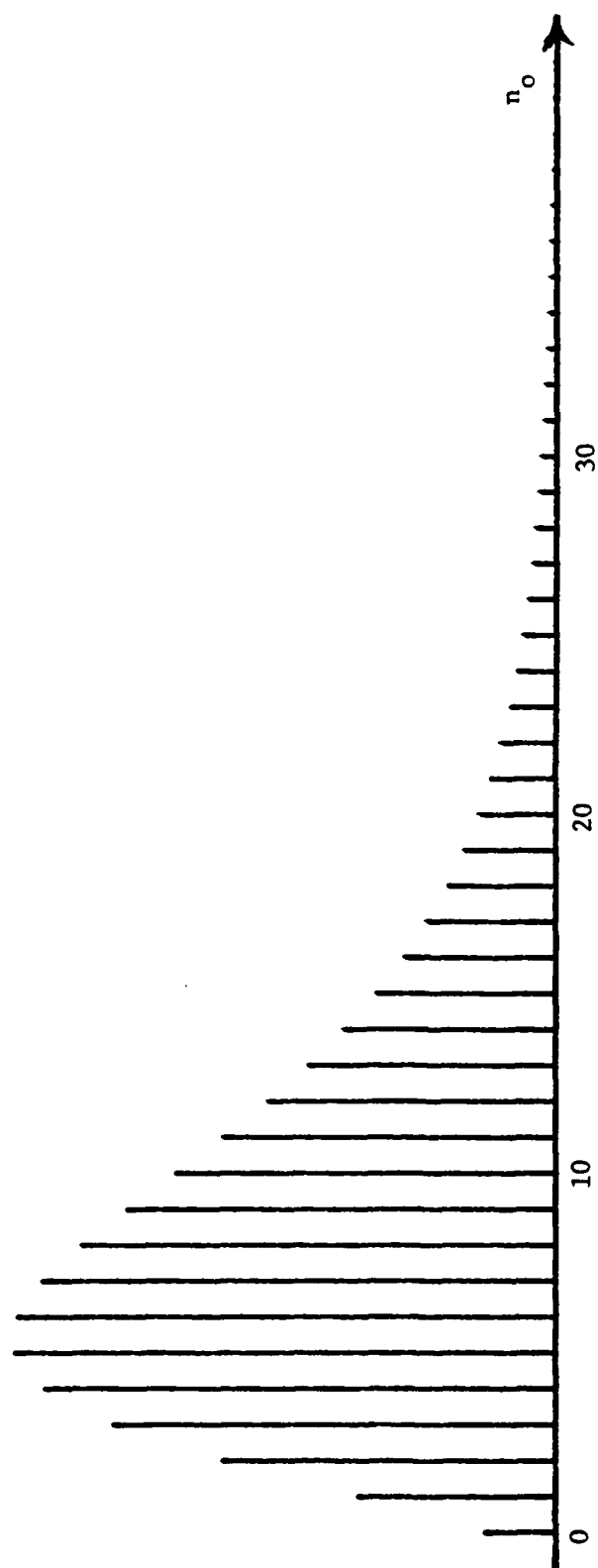


Figure 5. Predictive density of undetected errors at $t = 10$

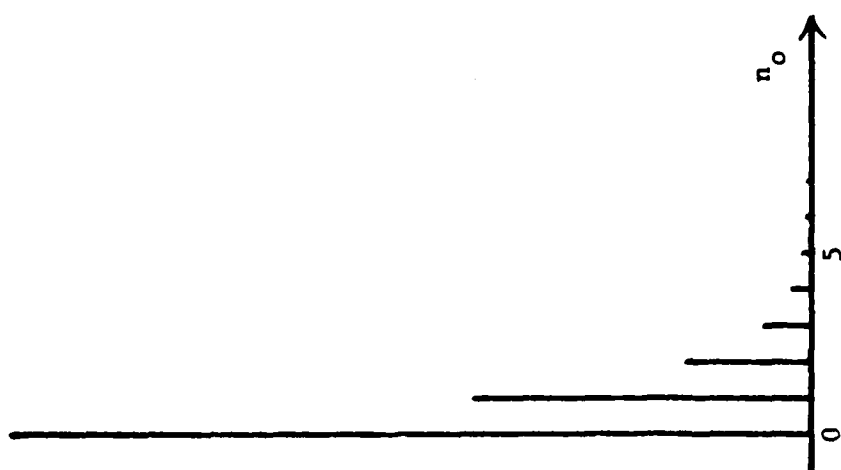


Figure 6. Predictive density of undetected errors at $t=70$

References

- [1] A. Goel and K. Okumoto, "Time-Dependent Error-Detection Rate Model for Software Reliability and Other Performance Measures," IEEE Transactions on Reliability, Vol. R-28, No. 3, pp. 206-11, 1979.
- [2] Z. Jelinski and P. B. Moranda, "Software Reliability Research," in Statistical Computer Performance Evaluation (W. Freiberger, ed.), Academic Press, NY, pp. 485-502, 1972.
- [3] W. S. Jewell, "A General Framework for Learning Curve Reliability Growth Models," Operations Research, Vol. 32, No. 3, pp. 547-58, May-June 1984.
- [4] W. S. Jewell, "Bayesian Estimation of Undetected Errors," ORC 83-11, Operations Research Center, University of California, Berkeley; to appear, Proceedings, 2nd International Meeting on Bayesian Statistics, Valencia, Spain, 6-10 September 1983.
- [5] W. Kremer, "Birth-Death and Bug Counting," IEEE Transactions on Reliability, Vol. R-32, No. 1, pp. 37-47, 1983.
- [6] N. Langburg and N. D. Singpurwalla, "Unification of Some Software Reliability Models via the Bayesian Approach," SIAM Journal of Scientific and Statistical Computing, Vol. 6, No. 3, 1985.
- [7] B. Littlewood and J. Verall, "A Bayesian Reliability Growth Model for Computer Software," Record of the IEEE Symposium on Computer Software Reliability, IEEE, NY, pp. 70-76, 1973.
- [8] B. Littlewood and J. Verall, (same title as [7]), Journal Royal Statistical Society, Vol. 22, No. 3, pp. 332-46, 1973.
- [9] B. Littlewood, "How to Measure Software Reliability and How Not To," IEEE Transactions on Reliability, Vol. R-28, No. 2, pp. 103-10, 1979.
- [10] B. Littlewood, "Stochastic Reliability-Growth: A Model for Fault-Removal in Computer-Programs and Hardware-Designs," IEEE Transactions on Reliability, Vol. R-30, No. 4, pp. 313-20, 1981.
- [11] R. J. Meinhold and N. D. Singpurwalla, "Bayesian Analysis of a Commonly Used Model for Describing Software Failures," The Statistician, Vol. 32, pp. 168-73, 1983.
- [12] P. B. Moranda, "An Error Detection Model for Application During Software Development," IEEE Transactions on Reliability, Vol. R-30, No. 4, pp. 309-13, 1981.

END

FILMED

9-85

DTIC